

## 機械学習手法を用いた音楽演奏時の呼吸の分析

五十嵐創<sup>†</sup> 尾崎知伸<sup>††</sup> 植野研<sup>†</sup> 古川康一<sup>†</sup>

<sup>†</sup>慶應義塾大学大学院 政策・メディア研究科

<sup>††</sup>慶應義塾大学 SFC 研究所

{soh,tozaki,ueno,furukawa}@sfc.keio.ac.jp

<http://bruch.sfc.keio.ac.jp/>

### 概要

本論文では、音楽演奏時の呼吸データに対する、決定木および帰納論理プログラミング (Inductive Logic Programming, ILP) を用いた分析実験について報告する。音楽演奏においては、楽曲の表情に応じて呼吸の柔軟なコントロールを行うことが、巧みな演奏を実現するための一つの重要な要素であると考えられる。本論文では生体センサから得られた呼吸データ、楽曲のフレーズ・和声構造などの楽譜情報、および奏法に関する情報を用いて、決定木および ILP により、音楽演奏時の演奏者の呼吸の状態に関するルールを抽出することを試みた。

## Analysis of Respiration during Musical Performances by Machine Learning

So Igarashi<sup>†</sup>, Tomonobu Ozaki<sup>††</sup>, Ken Ueno<sup>†</sup>, Koichi Furukawa<sup>†</sup>

<sup>†</sup>Graduate School of Media and Governance, Keio University

<sup>††</sup>Keio Research Institute at SFC

### Abstract

In this paper, we describe our analysis of respiration during musical performances by decision tree induction and Inductive Logic Programming (ILP). For good (skilled) musical performances, it can be regarded as one of the most important factors to control one's respiration in response to the aspect of music performed. It is, however, often difficult to explain how to do it clearly even for experts. This is a kind of skill clarification problem which is important in cognitive, biomechanical and other sciences. This paper takes a cello performance as an example. We measured respiration during a cello performance, and tried to extract rules of respiration from the data with musical background knowledge such as harmonic progression and bowing.

## 1 問題設定と背景

世の中にはいわゆる上手なとか、良いとされる音楽演奏がある。すべての演奏家はより上手く、よ

り良く演奏できるよう努力をする。一般的に、上手な演奏、良い演奏とそうでない演奏とは、演奏者自身、および聴取者の耳によって分類されるが、なぜその演奏が上手なのか、なぜ良いのかという

ことを明確に説明することは難しい。またさらに、どのようにすれば上手な、良い演奏ができるのか、というすべての演奏家および音楽愛好家にとって最も切実な問いに対しても、彼らを十分に満足させ得るような、具体的で科学的な答えはいまだ見つかっていない。しかし我々は経験を通じて、いわゆる上手な、良いとされる演奏の中には、何らかの法則や共通する特徴があることを感じ取ることができる。そのような何らかの法則や特徴を一種の制約と考え、熟練者とは、そのような制約を満たすことができるような演奏者であるということが出来る [Furukawa 00, Ueno 01]。ここで問題になるのは、そのような制約は、当の熟練者自身にとっても明確には意識されていない場合が多いということである。つまり、上手な、良い演奏をするためにはどのような制約を満たすことが必要なのか、熟練者自身でも明確に説明することができないことが多いのである。であるから、もしそのような制約を明示的に言語によって表現することができれば、それは多くの人々にとって非常に興味深い知識といえるだろう。

本研究では、具体的にそのような制約の一つとして、チェロ演奏時の呼吸の状態に着目する。音楽演奏における呼吸の役割としては、[中村 94, 長岡 99] などにより、演奏における「間」の取り方と呼吸との間に密接な関係があることが報告されている。さらに演奏のスキルという観点から言うと、経験的に、熟練者は音楽の表情に合わせた巧みな呼吸のコントロールを行っていると考えられる。生体センサを用いた実験から得られた呼吸のデータと、演奏される楽曲およびその奏法についての情報をもとに、決定木 [Quinlan 93] や帰納論理プログラミング (Inductive Logic Programming, ILP) [古川 01] などの機械学習手法によって、音楽演奏時における呼吸についてのルールを発見することが本研究の目的である。

## 2 データ獲得

データ獲得のための実験を行い、チェロ演奏時の呼吸の状態を計測した。被験者は演奏歴約 40 年の

男性アマチュア演奏家 1 名であり、タスクは Luigi Boccherini の Rondo 中の第 15 小節から第 38 小節までである (図 1)。分析の単純化のため、毎分 4 分音符=64 のテンポに設定したメトロノームに合わせて演奏してもらった。呼吸センサは胴体部巻きつけ式を用いた。

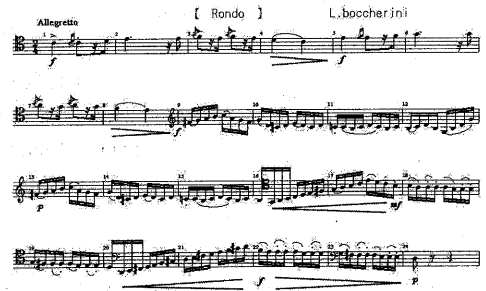


図 1: 実験のタスク、Boccherini の Rondo の一部

実験によって得られたデータにスムージングなどの前処理を施したものを図 2 に示す。これは呼吸曲線と呼ばれる、人間の呼吸の状態を電圧の変化に変換して表したグラフである。横軸は時間軸である。また、縦軸は「近似的に」呼吸の量を表すと考えることができる。グラフが下がっている部分は息を吸っている状態を表し、上がっている部分は吐いている状態を表す。ゆえに、例えばグラフが急に右下がりになっている部分は被験者が大きく急に息を吸う状態を表し、緩やかな傾きで徐々に上がっている部分はゆっくり息を吐いている状態を表す。グラフがほとんど水平で変化がない部分は、被験者があまり激しい呼吸をしていないことを示す。グラフが上に、あるいは下に凸になっている部分は、(呼→吸および吸→呼の)呼吸の変化点を表す。

図 2 の横軸方向の目盛は、1 目盛が 1 小節を表す。この曲は 4 分の 2 拍子であるため、1 小節は 2 拍からなる。また、図 2 中の矢印のついた目盛は、最初と最後が曲の開始と終わりを、真ん中がそこを境に曲調が情緒的なものから技巧的なものに変わることを示している。

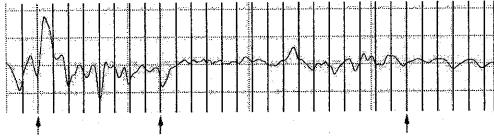


図 2: チェロ演奏時の演奏者の呼吸曲線

### 3 機械学習によるアプローチ

本研究では、前節で述べたような形で得られたデータに対して、機械学習によるアプローチを採用した。具体的には、ID3(C4.5) アルゴリズムに基づく決定木による分類、および ILP によるルール抽出を行った。決定木による分類は、命題論理上で帰納推論を行う枠組みであり、大量のデータの中に潜む法則や傾向を発見するデータマイニングの手法として、現在最も一般的なものの一つである。そして C4.5 は、情報利得比基準に基づく、代表的な決定木生成アルゴリズムである。また、ILP は述語論理上で帰納推論を行う枠組みであり、対象間の関係の記述能力に優れている。決定木による分類との比較においては、各事例の持つ属性のみならず、それらの背景知識を学習に際して利用できる点が特長である。

このような機械学習の手法を、音楽、あるいは人間の身体スキルの問題に適用した研究はいくつかなされている。C.Sammut らは、C4.5 によって航空機の操縦技術に関するルールを抽出し、フライトシミュレータを用いてその実証実験を行った [Sammut 92]。また、I.Bratko らは、同じく C4.5 を用いてクレーンの操縦技術の抽出を行った [Bratko 99]。また Widmer は、音楽演奏の分析に ILP を応用した [Widmer 96]。

決定木および ILP は、ともに論理に基づく分類学習器であるため、本研究で用いる生体データのような、時系列の連続データをそのままの形で扱うことは困難であり、前処理として適当にデータを離散化する必要がある。具体的には図 2 に示したような曲線として得られるデータをどのように区分け(セグメンテーション)し、一つの事例とみなすかを決めなければならない。この、データの

セグメンテーションが本研究における一つの重要な問題である。そしてそれに伴い、演奏される楽曲の楽譜情報(和声構造や、奏法など)を、属性および背景知識としてどのように与えるかということがもう一つの問題となる。

### 4 決定木による分類

以上の点をふまえ、まず初期実験として、決定木による分類学習を行った。実際には、Weka [Witten 00] の中で提供されている、C4.5 に基づく分類学習器 J48 を使用した。セグメンテーションについては楽譜上の 1 拍を 1 事例とし、それぞれについて呼吸の状態を呼吸曲線に基づき、吸う (inspiration)、吐く (expiration)、無呼吸 (no) の 3 種に分類し、学習の際のクラスとした。1 拍の中に呼吸の変化点を含むような事例に対してはクラスを与えなかった。このタスクは全 24 小節、48 拍からなる。そのうち最後の 1 拍は休符のみからなるため分析からは除外した。結果として 3 クラス 36 事例が得られた。

各事例が持つ属性については、その拍がいくつの音符からなるか (has\_notes)、音型 (shape)、フレーズの始まりかどうか (start\_phrase)、終わりかどうか (end\_phrase)、クレッシェンド (ディミヌエンド) しているかどうか (crescendo, diminuendo)、調性 (key)、和声機能 (harmonic\_function)、弓使い (bowing)、移弦の有無 (string\_transition)、左手のポジション移動の有無 (shift)、という楽譜情報および奏法に関する情報を与えた。これらのうち、has\_notes のみは数値属性で、あとはすべて名義属性である。

決定木による学習では、基本的に各事例間の関係は扱うことはできない。すなわちこの分析は、各拍における呼吸の状態が、その拍の持つ性質のみを使ってどのように説明されるか、という問題を解いていることになる。結果として、以下のような決定木が得られた。

```

has_notes <= 2
| start_phrase = yes: expiration (2.0)
| start_phrase = no: inspiration (9.0/2.0)
has_notes > 2
| harmonic_f = tonic
| | bowing = d: expiration (0.0)

```

```

| | bowing = u: no (1.0)
| | bowing = dd: expiration (0.0)
| | bowing = uu: expiration (0.0)
| | bowing = du: no (3.0/1.0)
| | bowing = dud: expiration (4.0/1.0)
| | bowing = udu: inspiration (2.0)
| | bowing = dudu
| | | start_phrase = yes: inspiration(2.0/1.0)
| | | start_phrase = no: expiration (2.0)
| harmonic_f = dominant: no (10.0/3.0)
| harmonic_f = subdominant: expiration (1.0)

```

各葉に割りふられた数字は、その葉に属する事例の数を表し、その右隣の数字はそのうち誤って分類されたものが何件あるかを示す。出力結果によれば、その拍に何個の音符があるかが、呼吸の状態を決定する一番大きな要因になっている。音符を2個以下しか持たない拍については、それがフレーズの最初の拍ならば息を吐き、最初の拍でなければ息を吸う、というルールが生成された。また、2個より多くの音符を持つ拍については、和声機能によって分類がなされた。和声機能がドミナントならば無呼吸であり、サブドミナントならば息を吐くというルールが生成された。また、和声機能がトニックのものに関しては、さらにボウイングによって分類がなされた。得られた決定木は以上のようなルールとして解釈されるが、この結果に対し、Leave-One-Outによる検定を行った結果、未知の事例に対する予測分類誤り率は約58%あった。結論として、ある拍の呼吸の状態を予測するのに、その拍自身の持つ性質のみでは十分ではないといえることができる。

次に、各事例の前後2事例の属性を、その事例の属性として与えて同様の学習を行った。この場合、学習の際に、対象となる事例(拍)の前後2事例(拍)に関する情報を扱えることになり、結果として、各事例の順序関係が部分的に表現されることになる。得られた木を以下に示す。

```

pp_has_notes <= 1: expiration (5.29/0.15)
pp_has_notes > 1
| has_notes <= 2: inspiration (7.71/0.85)
| has_notes > 2
| | pp_string_transition = yes: expiration(3.0)
| | pp_string_transition = no
| | | p_start_phrase = yes: expiration (2.0)
| | | p_start_phrase = no
| | | | pp_crescendo = yes: inspiration(2.0)
| | | | pp_crescendo = no: no (16.0/4.0)

```

木の中に現れた属性のうち、先頭に p\_ および pp\_ がつくものは、それぞれ1つ前および2つ前の事例(拍)の属性であることを表す。出力結果からは、2つ前の拍が1つしか音符を持たないならば息を吐く、というルールがまず読み取れる。全体的に、木の中には1つ前、あるいは2つ前の拍の属性が多くみられる。このことから、演奏中の呼吸が、直前にどのような音符を弾いたかということに強く影響を受けていることが予想される。しかし、Leave-One-Outによる検定を行った結果、未知の事例に対する予測分類誤り率は約69%と非常に高く、今回の結果だけでは、この予想はまだ確かなものとは言えないであろう。

## 5 ILP によるルール抽出

前節での決定木による分析の結果、各拍における呼吸の状態を予測するのに、その拍自身の持つ性質のみでは十分ではないことが示唆された。また、部分的に前後の事例に関する情報を用いても、あまり精度の高い予測は行われなことがわかった。そのため、今度はILPによるルール抽出を行った。データは同一のものだが、ILPに用いるため、各拍の属性を、以下のような述語形式に変換して表現した。そして、nextto という述語を新しく定義することにより、先ほどは直接扱うことができなかった事例間の関係、すなわち楽譜上における各拍の順序関係を表現した。実際の記述例を以下に示す。

```

class(beat39,inspiration).
‘‘第39拍では息を吸う’’
has_notes(beat39,4).
‘‘第39拍は4つの音符からなる’’
shape(beat39,ascending).
‘‘第39拍は上昇音型である’’
start_phrase(beat39).
‘‘第39拍はフレーズの最初の拍である’’
string_transition(beat39).
‘‘第39拍で移弦を行う’’
shift(beat39).
‘‘第39拍で左手のポジション移動を行う’’
crescendo(beat39).
‘‘第39拍はクレッシェンドである’’
key(beat39,d_major).
‘‘第39拍は二長調である’’

```

```

harmonic_function(beat39,tonic).
‘第39拍の和声機能はトニックである’
bowing(beat39,dudu).
‘第39拍の弓使いはダウン、アップ、ダウン、アップである’
nextto(beat38,beat39).
‘第38拍の次が第39拍である’
nextto(beat39,beat40).
‘第39拍の次が第40拍である’

```

以上のようなデータセットに対し、代表的な ILP システムの一つである Progol [Muggleton 95] による実験を行った。Progol では複数の仮説を推論するために、集合被覆アルゴリズム (cover set algorithm) を採用しているため、学習の際に事例を与える順序が、得られる結果に大きく影響をおよぼす可能性がある。そのため、事例を与える順番のみを無作為に変化させた5種類のデータセットを用意し、それぞれについて実験を行った。結果として、以下の3つのルールが5種類のデータセットすべてから得られた。

```

class(A,inspiration) :-
    nextto(A,B), bowing(B,du), shift(B).
‘次の拍の弓使いがダウン、アップでポジション移動を含むならば息を吸う’

```

```

class(A,expiration) :-
    nextto(B,A), bowing(B,dudu).
‘前の拍の弓使いがダウン、アップ、ダウン、アップならば息を吐く’

```

```

class(A,no) :-
    shape(A,d_to_a),
    harmonic_function(A,dominant).
‘下降して上昇する音型で、和声機能がドミナントならば無呼吸である’

```

また、次のルールが4種類のデータセットから共通して得られた。

```

class(A,expiration) :-
    nextto(B,A), nextto(C,B), has_note(C,1).
‘前の前の拍が1つの音符からなるなら息を吐く’

```

このルールは、決定木による2度目の分析でも得られている。以上の4つのルールは、このデー

タセットから得られる妥当なルールであると考えられる。ここまで得られたルールを見てみると、吸気 (inspiration) に関するルールはその拍より前の拍の状態に依存するのに対し、呼気 (expiration) の場合はその拍より後の、無呼吸 (no) の場合はその拍自身の状態に依存することが分かる。このことから、息を吸うという行為は、これからする演奏のための、一種の準備動作ではないかと考えることができる。この予想は、次に示す明快なルールによっても支持される。

```

class(A,inspiration) :- end_phrase(A).
‘フレーズの最後ならば息を吸う’

```

このルールは3種類のデータセットから得られた。フレーズの最後で息を吸うというのはおかしな感じがするかもしれないが、これも室内楽奏者達が同時に息を吸うことによって次の音の出のタイミングを合わせるように、新しく始まるフレーズのための準備動作の一つと解釈することもできる。これらの他に、2つのデータセットから得られたルールが1つ、1つのデータセットのみから得られたルールが3つあった。

Leave-One-Out による検定を行ったところ、各データセットの精度は約65%であった。また、すべてのデータセットにおいて、一般化されない事例が、全事例の約4割にあたる15事例ほどあった。先の図2を見ると、これらの事例(拍)は、(吸っているにせよ吐いているにせよ)比較的呼吸の変化が小さい部分であることが分かった。呼吸は、人間が生きている限り、どのような状況であれ行われる現象なので、これら一般化されなかった事例というのは、特に音楽演奏と関係なく、単に生理的な必要性からなされたものである可能性がある。

## 6 今後の課題

今後の課題としては、まず全般的な問題として、呼吸データのセグメンテーションに改良の余地がある。今回は単純に楽譜上の1拍を1事例として全体を等分したが、そうすると1拍の中に複数のクラスが含まれる(1拍の間に呼息と吸息が両方含

まれる)という事態が起こってしまうため、より適切なセグメンテーションを行う必要がある。また、今回の分析では呼吸の状態について、特に吐いているか吸っているかという視点で分類を行ったが、もう一つ、演奏中の息継ぎ、すなわち呼息から吸息への変化点に注目するという方法も当然考えられる。さらに、個々の事例にクラスを割り当てる際、先の実験のように、単純に吸っているか吐いているか無呼吸であるかの3つに分類してしまうと、呼吸の深さという要素が失われてしまい、先ほどのように、本当に音楽的な理由でコントロールされて行われている呼吸と、単に生理的な必要性から行われているであろう呼吸とを混同してしまうことになる。今後は以上のような問題を視野に入れ、より良いセグメンテーションの方法を考え出す必要がある。

次に機械学習の視点から言えば、ILPの背景知識として与える音楽知識を洗練することが考えられる。現時点で扱っている情報はかなり限定されているが、より抽象度の高い概念、複雑な音楽構造などを背景知識として与えることにより、人間にとってより理解しやすい、発見的価値のあるルールを抽出できる可能性がある。

今後は、以上の点をふまえた上で、他のILPシステムでの比較実験、新データを用いた追実験などの拡張を行う。

## 参考文献

- [Bratko 99] Bratko, I. and Urbancic, T.: Control Skill, Machine Learning and Hand-crafting in Controller Design, In Furukawa, K., Michie, D., and Muggleton, S. eds., *Machine Intelligence 15*, Oxford Press, 1999.
- [Furukawa 00] Furukawa, K., Ueno, K., and Bain, M.: Motor Skill as Dynamic Constraint Satisfaction, *Proceedings of the 17th International Workshop on Machine Intelligence*, pp.25-28, 2000.
- [古川 01] 古川康一, 尾崎知伸, 植野研: 帰納論理プログラミング, 共立出版, 2001.
- [Muggleton 95] Muggleton, S.: Inverse Entailment and Progol, *New Generation Computing*, Vol.13, No.3-4, pp.245-286, 1995.
- [中村 94] 中村敏枝: 音楽における「間」と呼吸について, 日本音響学会音楽音響研究会資料 MA94-16, 1994
- [長岡 99] 長岡千賀, 小森政嗣, 中村敏枝: 音楽における「間」の長さの判断と呼吸の関連, ヒューマンインターフェースシンポジウム'99, pp.399-404, 1999.
- [Quinlan 93] Quinlan, J.R.: *C4.5 Programs for Machine Learning*, Morgan Kaufmann, 1993.
- [Sammut 92] Sammut, C., Hurst, S., Kedzier, D., and Michie, D.: Learning to Fly, In Sleeman, D. and Edwards, P. eds., *Proceedings of the Ninth International Workshop on Machine Learning*, pp.385-393, Morgan Kaufmann, 1992.
- [Ueno 01] Ueno, K., Furukawa, K., and Bain, M.: Motor Skill as Dynamic Constraint Satisfaction, *Electric Transaction of Artificial Intelligence (ETAI)*, Linkoping University Electronic Press, 2001.
- [Widmer 96] Widmer, G.: Learning Expressive Performance: The Structure-Level Approach, *Journal of New Music Research*, Vol.25, No.2, pp. 179-205, 1996.
- [Witten 00] Witten, I.H. and Frank, E.: *Data Mining*, Morgan Kaufmann, 2000.